

Db 421 GTDFNTMTKIFSDKASISNEDNMFKITNTLVKGFADPVRSTINLSGPGVYTGAYQD 480  
 QY 481 LLMFKDNTNHLLEPELRFEPKTKISQITQETISLASFNOARAKSOFEEYKKGYPE 540  
 Db 481 LLMFKDNTNHLLEPELRFEPKTKISQITQETISLASFNOARAKSOFEEYKKGYPE 540  
 QY 541 GALGEDNDLFAQNTVLDKDYVSKKILSSMKTNRKNEYIHYIVOLQGDKISYEASCNLFK 600  
 Db 541 GALGEDNDLFAQNTVLDKDYVSKKILSSMKTNRKNEYIHYIVOLQGDKISYEASCNLFK 600  
 QY 601 DPTSSILYQKNIEGSETAYYYYVADAEIKEDIRYPIQISNKRNIKLTFIGHGKSEFNT 660  
 Db 601 DPTSSILYQKNIEGSETAYYYYVADAEIKEDIRYPIQISNKRNIKLTFIGHGKSEFNT 660  
 QY 661 DTFANLDVSLSEIETILNLAADISPKYIEINLGCNMFYSIYAEETYPGKLLKIK 720  
 Db 661 DTFANLDVSLSEIETILNLAADISPKYIEINLGCNMFYSIYAEETYPGKLLKIK 720  
 QY 721 DRYSELMPSISQDSITVSANQYEVRIEKGKREILDSHGWINKERSIINDISSKEYISF 780  
 Db 721 DRYSELMPSISQDSITVSANQYEVRIEKGKREILDSHGWINKERSIINDISSKEYISF 780  
 QY 781 NPENKIIYKSKYLHELSTLLOEIRNANSSDIDLEKKVMTLCEINVASNIDRQIVEGR 840  
 Db 781 NPENKIIYKSKYLHELSTLLOEIRNANSSDIDLEKKVMTLCEINVASNIDRQIVEGR 840  
 QY 841 IEAKNLTSDSINYIKNEFKLIESDSLYDLKHQGLDSDHFISSDKTENGFRIRF 900  
 Db 841 IEAKNLTSDSINYIKNEFKLIESDSLYDLKHQGLDSDHFISSDKTENGFRIRF 900  
 QY 901 INKETGNSIFETEKEIFSEYATHISKEISNKTDTFDNVNGKLVKVNDAHAHVNTLN 960  
 Db 901 INKETGNSIFETEKEIFSEYATHISKEISNKTDTFDNVNGKLVKVNDAHAHVNTLN 960  
 QY 961 SAFFIOSLIEYNTKESLNSLVAMKVQVYAGLFSTGLNTITDASKVVELSTALDETID 1020  
 Db 961 SAFFIOSLIEYNTKESLNSLVAMKVQVYAGLFSTGLNTITDASKVVELSTALDETID 1020

## RESULT 2

S10317  
 toxin B - Clostridium difficile  
 C:Species: Clostridium difficile  
 C:Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 15-Oct-1999  
 C:Accession: S10317; S21894; S22434  
 R:Barroso, L.A.; Wang, S.Z.; Phelps, C.J.; Johnson, J.L.; Wilkins, T.D.  
 Nucleic Acids Res. 18, 4004, 1990  
 A:Title: Nucleotide sequence of Clostridium difficile toxin B gene.  
 A:Reference number: S10317; MUID:90326540  
 A:Accession: S10317  
 A:Status: translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-2366 <BAR>  
 A:Cross-references: EMBL:X5138; NID:940442; PIDN:CAA37298.1; PID:940443  
 R:Eigen-Streiber, C.; Laufenberg-Feldmann, R.; Sartingen, S.; Schulze, J.; Sauerborn, A.  
 submitted to the EMBL Data Library, July 1991  
 A:Description: Comparative analysis of Clostridium difficile toxins A and B.  
 A:Reference number: S21894  
 A:Accession: S21894  
 A:Molecule type: DNA  
 A:Residues: 1271-2366 <EIC>  
 A:Cross-references: EMBL:X60984; NID:940445; PIDN:CAA43299.1; PID:940446  
 R:Eigen-Streiber, C.; Laufenberg-Feldmann, R.; Sartingen, S.; Schulze, J.; Sauerborn, A.  
 Mol. Genet. 233, 260-268, 1992  
 A:Title: Comparative sequence analysis of the Clostridium difficile toxins A and B.  
 A:Reference number: S22434; MUID:92293124  
 A:Accession: S22434  
 A:Molecule type: DNA  
 A:Residues: 1791-2366 <VON>  
 A:Cross-references: EMBL:X60984  
 C:Genetics:  
 A:Gene: toxB

C:Superfamily: cpl repeat homology  
 C:Keywords: cytotoxin

Query Match 78.7%; Score 4086; DB 2; Length 2366;  
 Best Local Similarity 77.5%; Pred. No. 4.9e-146;  
 Matches 790; Conservative 105; Mismatches 125; Indels 0; Gaps 0;

QY 1 MSLVNRKQLEKMANVRFQDEYVAILDALEEHNMSENTEVYKLVKLDINSUTDIYI 60  
 Db 1 MSLVNRKQLEKMANVRFQDEYVAILDALEEHNMSENTEVYKLVKLDINSUTDIYI 60  
 QY 61 NYKSGRNKALKKFEYLTMEVLEKNNSLTPVEKNLHFIIWGGINDTAINYNQWKD 120  
 Db 61 NYKSGRNKALKKFEYLTMEVLEKNNSLTPVEKNLHFIIWGGINDTAINYNQWKD 120  
 QY 121 VNSDITVRFYDSNAPLNTLAKTIVESANTNTLESFRENLDPEFYKPKRMEIY 180  
 Db 121 VNSDITVRFYDSNAPLNTLAKTIVESANTNTLESFRENLDPEFYKPKRMEIY 180  
 QY 181 DKQKHFIDYKSOIBENPEFIIDNIKYLSNEYSKOLEALNKYIEESLNKITANGNDI 240  
 Db 181 DKQKHFIDYKSOIBENPEFIIDNIKYLSNEYSKOLEALNKYIEESLNKITANGNDI 240  
 QY 241 RNLEKFADEDLVRVNLQELVERNNLAAASDLIRISMLKEDGGYLDVLDILPGIOPDLFS 300  
 Db 241 RNLEKFADEDLVRVNLQELVERNNLAAASDLIRISMLKEDGGYLDVLDILPGIOPDLFS 300  
 QY 301 INKPSITNTSWEMIKLEAIKMYKEYIPGYTSKNFMDLDEEVQSFESALSCKSKSEIF 360  
 Db 301 INKPSITNTSWEMIKLEAIKMYKEYIPGYTSKNFMDLDEEVQSFESALSCKSKSEIF 360  
 QY 361 LPDIDIKYPLEVIAFANNSVINOALISLKDSCSDIVINOIKNRYKILNDNLNPSINE 420  
 Db 361 LPDIDIKYPLEVIAFANNSVINOALISLKDSCSDIVINOIKNRYKILNDNLNPSINE 420  
 QY 421 GTDNTTKIIFSDKLASISNEDNMFMKINYLKVGFPADVRSTINLSGPGVYTGAYQD 480  
 Db 421 GTDNTTKIIFSDKLASISNEDNMFMKINYLKVGFPADVRSTINLSGPGVYTGAYQD 480  
 QY 481 LLMFKDNTNHLLEPELRFEPKTKISQITQETISLASFNOARAKSOFEEYKKGYPE 540  
 Db 481 LLMFKDNTNHLLEPELRFEPKTKISQITQETISLASFNOARAKSOFEEYKKGYPE 540  
 QY 541 GALGEDNDLFAQNTVLDKDYVSKKILSSMKTNRKNEYIHYIVOLQGDKISYEASCNLFK 600  
 Db 541 GALGEDNDLFAQNTVLDKDYVSKKILSSMKTNRKNEYIHYIVOLQGDKISYEASCNLFK 600  
 QY 601 DPTSSILYQKNIEGSETAYYYYVADAEIKEDIRYPIQISNKRNIKLTFIGHGKSEFNT 660  
 Db 601 DPTSSILYQKNIEGSETAYYYYVADAEIKEDIRYPIQISNKRNIKLTFIGHGKSEFNT 660  
 QY 661 DTFANLDVSLSEIETILNLAADISPKYIEINLGCNMFYSIYAEETYPGKLLKIK 720  
 Db 661 DTFANLDVSLSEIETILNLAADISPKYIEINLGCNMFYSIYAEETYPGKLLKIK 720  
 QY 721 DRYSELMPSISQDSITVSANQYEVRIEKGKREILDSHGWINKERSIINDISSKEYISF 780  
 Db 721 DRYSELMPSISQDSITVSANQYEVRIEKGKREILDSHGWINKERSIINDISSKEYISF 780  
 QY 781 NPENKIIYKSKYLHELSTLLOEIRNANSSDIDLEKKVMTLCEINVASNIDRQIVEGR 840  
 Db 781 NPENKIIYKSKYLHELSTLLOEIRNANSSDIDLEKKVMTLCEINVASNIDRQIVEGR 840  
 QY 841 IEAKNLTSDSINYIKNEFKLIESDSLYDLKHQGLDSDHFISSDKTENGFRIRF 900  
 Db 841 IEAKNLTSDSINYIKNEFKLIESDSLYDLKHQGLDSDHFISSDKTENGFRIRF 900  
 QY 901 INKETGNSIFETEKEIFSEYATHISKEISNKTDTFDNVNGKLVKVNDAHAHVNTLN 960  
 Db 901 INKETGNSIFETEKEIFSEYATHISKEISNKTDTFDNVNGKLVKVNDAHAHVNTLN 960  
 QY 961 SAFFIOSLIEYNTKESLNSLVAMKVQVYAGLFSTGLNTITDASKVVELSTALDETID 1020  
 Db 961 SAFFIOSLIEYNTKESLNSLVAMKVQVYAGLFSTGLNTITDASKVVELSTALDETID 1020

Sat Aug